ON THE LASSERRE HIERARCHY OF SEMIDEFINITE PROGRAMMING RELAXATIONS OF CONVEX POLYNOMIAL OPTIMIZATION PROBLEMS

ETIENNE DE KLERK* AND MONIQUE LAURENT†

Abstract. The Lasserre hierarchy of semidefinite programming approximations to convex polynomial optimization problems is known to converge finitely under some assumptions. [J.B. Lasserre. Convexity in semialgebraic geometry and polynomial optimization. *SIAM J. Optim.* **19**, 1995–2014, 2009.] We give a new proof of the finite convergence property, under weaker assumptions than were known before. In addition, we show that — under the assumptions for finite convergence — the number of steps needed for convergence depends on more than the input size of the problem.

Key words. convex polynomial optimization, sum of squares of polynomials, positivstellensatz, semidefinite programming.

AMS subject classifications. 90C60, 90C56, 90C26

1. Polynomial optimization and the Lasserre hierarchy.

We consider the polynomial optimization problem

$$p_{\min} = \min_{x\in\mathbb{R}^n} \{p_0(x) \;:\; p_i(x) \geq 0 \; (i = 1, \ldots, m)\}, \tag{1.1}$$

where each $p_i : \mathbb{R}^n \to \mathbb{R}$ $(i = 0, \ldots, m)$ is a polynomial. We denote the highest total degree of the polynomials $p_0, \ldots, p_m$ by d. We partition the index set $\{1, \ldots, m\} =: \mathcal{I}_l \cup \mathcal{I}_n$ to differentiate between (affine) linear and nonlinear constraints, where $\mathcal{I}_l$ consists of the indices i for which p_i is an affine or linear polynomial.

We denote the polynomials with real coefficients in the variables x by $\mathbb{R}[x]$. The subset of $\mathbb{R}[x]$ consisting of the sums of squares of polynomials is denoted by Σ^2.

The feasible set of problem (1.1) is denoted by $\mathcal{F}$, i.e:

$$\mathcal{F} := \{x \in \mathbb{R}^n \mid p_i(x) \geq 0 \; (i = 1, \ldots, m)\}. \tag{1.2}$$

The quadratic module generated by the polynomials p_i $(i = 1, \ldots, m)$ is defined as:

$$\mathcal{M}(p_1, \ldots, p_m) := \left\{ \sigma_0 + \sum_{i=1}^{m} \sigma_i p_i \;\middle|\; \sigma_i \in \Sigma^2 \; (i = 0, \ldots, m) \right\}. \tag{1.3}$$

The quadratic module $\mathcal{M}(p_1, \ldots, p_m)$ is called *Archimedean* if

$$\exists R > 0 \;:\; R^2 - \|x\|^2 \in \mathcal{M}(p_1, \ldots, p_m).$$

Throughout the paper, we will assume that $\mathcal{M}(p_1, \ldots, p_m)$ is Archimedean. Note that this assumption implies that $\mathcal{F}$ is compact (since it is contained in the ball $B(0, R) := \{x : \|x\| \leq R\}$). Moreover, we may assume without loss of generality that $\|x\| < R$ for all $x \in \mathcal{F}$, since $\bar{R}^2 - \|x\|^2 \in \mathcal{M}(p_1, \ldots, p_m)$ for all $\bar{R} \geq R$.

*Tilburg University, Department of Econometrics and Operations Research, 5000 LE Tilburg, Netherlands. Email: E.deKlerk@uvt.nl

†Centrum Wiskunde & Informatica (CWI), Science Park 123, 1098 XG Amsterdam, and Tilburg University, Department of Econometrics and Operations Research, 5000 LE Tilburg, Netherlands. Email: M.Laurent@cwi.nl

The truncated quadratic module of degree $2t$, denoted as $\mathcal{M}_t(p_1,\ldots,p_m)$, is defined as the subset of $\mathcal{M}(p_1,\ldots,p_m)$ where the sum of squares polynomials $\sigma_0,\ldots,\sigma_m$ meet the additional conditions:

$$\deg(\sigma_0) \le 2t,\ \deg(\sigma_i p_i) \le 2t \quad (i=1,\ldots,m). \tag{1.4}$$

Lasserre [11] introduced the following hierarchy of approximations to $p_{\min}$:

$$\rho_t := \max\{\lambda \mid p_0 - \lambda \in \mathcal{M}_t(p_1,\ldots,p_m)\} \tag{1.5}$$

and showed that, under the Archimedean assumption, $\lim_{t\to\infty} \rho_t = p_{\min}$. Moreover, for each fixed t, ρ_t may be computed as the optimal value of a semidefinite program. In particular, this may be done in polynomial time to any fixed accuracy.

Lasserre [12] recently showed that the hierarchy of approximations (1.5) exhibits finite convergence for certain classes of convex polynomial optimization problems (Theorem 3.4 in [12]).

Outline and scope of the paper. The purpose of our paper is to gain more insight into the convergence behavior of the Lasserre hierarchy. We will prove finite convergence for polynomial optimization problems that meet weaker conditions than those of Theorem 3.4 in [12]; see Theorem 3.2. In particular, finite convergence holds for convex instances (assuming positivity of the Hessian of the objective at the global minimizer); see Corollary 3.3. This should be seen in the light of the result of Helton and Nie [8] showing that regularity and positive curvature at the boundary of a convex semialgebraic set are sufficient to allow semidefinite programming representability via a finite number of liftings.

Moreover, we will construct some ‘bad’ examples where — under the assumptions for finite convergence — the number of steps for convergence cannot be bounded in terms of the problem data; see Theorem 4.2. In analogy, recall that no results are known that give degree bounds for sums of squares certificates of positivity of a polynomial depending only on its degree and number of variables; known bounds depend also on the minimum taken by the polynomial (see e.g. [16]).

2. Preliminaries on Lagrange multipliers and saddle points. We start by reviewing some well-known results in nonlinear programming (NLP). Proofs of all the results in this section may by found e.g. in the textbook by Bertsekas [4]. Consider the general NLP problem

$$f^* =: \inf_{x\in\mathcal{C}} \{f(x) \mid g_j(x) \le 0 \ (j=1,\ldots,m)\}, \tag{2.1}$$

where $f, g_1,\ldots,g_m$ are continuous functions defined on a set $\mathcal{C} \subset \mathbb{R}^n$.

Definition 2.1. *[4, Definition 5.1.1] The Lagrangian function of the optimization problem (2.1) is defined as:*

$$L(x,y) := f(x) + \sum_{j=1}^m y_j g_j(x), \qquad (x\in\mathcal{C},\ y\in\mathbb{R}^m_+).$$

A vector $\bar{y} \in \mathbb{R}^m_+$ is called a vector of Lagrange multipliers of (2.1) if

$$f^* = \inf_{x\in\mathcal{C}} L(x,\bar{y}).$$

A pair $(\bar{x},\bar{y}) \in \mathcal{C}\times\mathbb{R}^m_+$ is called a saddle point of the Lagrangian function if

$$L(\bar{x},y) \le L(\bar{x},\bar{y}) \le L(x,\bar{y}) \quad \forall\ (x,y)\in\mathcal{C}\times\mathbb{R}^m_+.$$

Lagrange multipliers, global minimizers and saddle points are related through the saddle point theorem.

THEOREM 2.2 (Saddle Point Theorem). *[4, Proposition 5.1.6] The pair* $(\bar{x}, \bar{y}) \in \mathcal{C} \times \mathbb{R}^m_+$ *is a saddle point if and only if* $\bar{x}$ *is a global minimizer of (2.1) and* $\bar{y}$ *is a vector of Lagrange multipliers of (2.1).*

Finally, by the convex Farkas lemma, the existence of Lagrange multipliers is assured in the convex case, under a suitable constraint qualification.

THEOREM 2.3 (Convex Farkas Lemma). *[4, Proposition 5.3.1] Assume that the functions* $f, g_1, \ldots, g_m$ *are convex, that the set* $\mathcal{C}$ *is a convex set, and that* f^* *is finite. Assume moreover, that there exists an* $x^* \in relint(\mathcal{C})$ *such that* $g_j(x^*) < 0$ *for all* j *with* g_j *non-linear, and* $g_j(x^*) \leq 0$ *for all* j *with* g_j *linear (the Slater constraint qualification). Then problem (2.1) has a vector of Lagrange multipliers.*

3. Finite convergence of the Lasserre hierarchy.

3.1. The general result. The aim in this section is to give a proof of the finite convergence result by Lasserre (Theorem 3.4 in [12]) under weaker assumptions. A key lemma that we will need is the following *Positivstellensatz* by Scheiderer [19].

PROPOSITION 3.1 (Example 3.18 in [19]). *Let* $p \in \mathbb{R}[x]$ *be a polynomial for which the level set*

$$\mathcal{K} := \{x \in \mathbb{R}^n \mid p(x) \geq 0\}$$

is compact. Let $q \in \mathbb{R}[x]$ *be nonnegative on* $\mathcal{K}$*. Assume that the following conditions hold:*

1. *q has only finitely many zeros in $\mathcal{K}$, each lying in the interior of K.*
2. *the Hessian $\nabla^2 q$ is positive definite at each of these zeroes.*

Then $q = \sigma_0 + p\sigma_1$ *for some* $\sigma_0, \sigma_1 \in \Sigma^2$.

We now prove the main result of this section, namely that the Lasserre SDP hierarchy has finite convergence for problem (1.1) under suitable assumptions.

THEOREM 3.2. *Consider the polynomial optimization problem (1.1), with Lagrangian function*

$$L(x, y) := p_0(x) - \sum_{j=1}^{m} y_j p_j(x) \quad x \in \mathcal{C},\ y \in \mathbb{R}^m_+,$$

where $\mathcal{C} = B(0, R)$ *so that* $\mathcal{F} \subset \mathcal{C}$*. Assume:*

1. *The quadratic module $\mathcal{M}(p_1, \ldots, p_m)$ is Archimedean;*
2. *There are finitely many global minimizers and at least one saddle point of L;*
3. *If $(\bar{x}, \bar{y})$ is a saddle point of L, then $\nabla^2_x L(\bar{x}, \bar{y}) \succ 0$.*

Then one has finite convergence of the Lasserre hierarchy, i.e:

$$p_0 - p_{\min} \in \mathcal{M}(p_1, \ldots, p_m).$$

Proof. Let $(\bar{x}, \bar{y})$ be a saddle point of problem (1.1). We first show that the function

$$q(x) := L(x, \bar{y}) - p_{\min} \equiv p_0(x) - p_{\min} - \sum_{j=1}^{m} \bar{y}_j p_j(x) \tag{3.1}$$

has finitely many roots in $\mathcal{C}$ and all these roots lie in the interior of $\mathcal{C}$. Indeed, every root x^* of q corresponds to a saddle point $(x^*, \bar{y})$, by construction, and is therefore also a global minimizer of problem (1.1), by the saddle point theorem. Consequently, every root x^* lies in the interior of $\mathcal{C}$ since we have assumed that $\|x\| < R$ for all $x \in \mathcal{F}$.

We may now apply Proposition 3.1 with $p(x) := R^2 - \|x\|^2$ and q as defined in (3.1), to conclude that

$$p_0(x) - p_{\min} = \sum_{j=1}^{m} \bar{y}_j p_j(x) + \sigma_0(x) + \sigma_1(x)(R^2 - \|x\|^2)$$

for some $\sigma_0, \sigma_1 \in \Sigma^2$. Since $R^2 - \|x\|^2 \in \mathcal{M}(p_1, \ldots, p_m)$ by assumption, we obtain the required result. □

3.2. The convex case. In the convex case, we have the following corollary.

COROLLARY 3.3. *Consider problem (1.1) under the following assumptions:*

1. *The polynomials $p_0, -p_1, \ldots, -p_m$ are convex;*
2. *The Slater condition holds:*

$$\exists x_0 \in \mathbb{R}^n \ : \ p_i(x_0) > 0 \textit{ for } i \in \mathcal{I}_n \textit{ and } p_i(x_0) \geq 0 \textit{ for } i \in \mathcal{I}_l.$$

3. *The quadratic module $\mathcal{M}(p_1, \ldots, p_m)$ is Archimedean:*

$$\exists R > 0 \ : \ R^2 - \|x\|^2 \in \mathcal{M}(p_1, \ldots, p_m).$$

4. *$\nabla^2 p_0(x^*) \succ 0$ (i.e. the Hessian of p_0 at x^* is positive definite) if x^* is a minimizer of (1.1).*

Then one has finite convergence of the Lasserre hierarchy, i.e:

$$p_0 - p_{\min} \in \mathcal{M}(p_1, \ldots, p_m).$$

Proof. The required result follows from Theorems 2.3 and 3.2. □

REMARK 3.1. *The fourth assumption in Corollary 3.3 implies that the minimizer of (1.1) is unique. It is a weaker assumption than the corresponding assumption in Theorem 3.4 of Lasserre [12] which requires that $\nabla^2 f(x) \succ 0 \ \forall \ x \in \mathcal{F}$. For example, consider the problem*

$$\min_{x \in [-1,1]} x^4 + 2x.$$

Here the Hessian is not positive definite at $x = 0$, but it is positive definite at the global minimizer $x^* = -2^{-1/3}$.

REMARK 3.2. *Note that Corollary 3.3 remains valid under different constraint qualifications. For instance, instead of assuming the existence of a Slater point (as in the second condition of Corollary 3.3), we may require the Mangasarian-Fromovitz constraint qualification:*

$$\exists w \in \mathbb{R}^n \ w^T \nabla p_i(x^*) > 0 \ \forall i \in J^*, \tag{3.2}$$

where x^* is a minimizer of (1.1) and $J^* = \{i \in \{1, \ldots, m\} \mid p_i(x^*) = 0\}$ is the set of indices corresponding to the active constraints at x^*. Indeed, under (3.2), there exist *multipliers $\bar{y}_j \geq 0$ for which $\nabla p_0(x^*) - \sum_j \bar{y}_j \nabla p_j(x^*) = 0$ and $\bar{y}_j p_j(x^*) = 0$ for all j (see e.g. [17, §12.6]).*

As in the proof of Theorem 3.2, consider the polynomial $q := p_0 - p_{\min} - \sum_j \bar{y}_j p_j$. As q is convex and $\nabla q(x^*) = 0$, x^* is a global minimizer of q over $\mathbb{R}^n$ and thus $q \geq q(x^*) = 0$ on $\mathbb{R}^n$. We can now proceed as in the rest of the proof of Theorem 3.2.

REMARK 3.3. *The assumption that the Hessian of p_0 should be positive definite at the minimizer cannot be omitted in Corollary 3.3.*

To see this, consider the problem

$$p_{\min} = \min_{x \in \mathbb{R}^n} \{p_0(x) \ : \ 1 - \|x\|^2 \geq 0\}, \tag{3.3}$$

where p_0 is a convex form (i.e. homogeneous polynomial) of degree at least 4 that is not a sum of squares.

Then $p_{\min} = 0$. Indeed, convex n-variate forms are necessarily nonnegative on $\mathbb{R}^n$, since their gradients vanish at zero[1]*. On the other hand, they are not always sums of squares, as was shown by Blekherman [5].*[2]

By construction, problem (3.3) satisfies all the assumptions of Corollary 3.3, except for the positive definiteness of the Hessian at the minimizer.

Assume we have finite convergence of the Lasserre hierarchy for problem (3.3), i.e.

$$p_0 \in \Sigma^2 + (1 - \|x\|^2)\Sigma^2.$$

By Proposition 4 in De Klerk, Laurent and Parrilo [10], a form belongs to the set $\Sigma^2 + (1 - \|x\|^2)\Sigma^2$ if and only if it is a sum of squares. This contradicts our assumption that $p_0 \notin \Sigma^2$.

REMARK 3.4. *Stronger finite convergence results than in Corollary 3.3 are known if the polynomials $p_0, -p_1, \ldots, -p_m$ are:*

1. *convex quadratic polynomials; here the Lasserre hierarchy is exact at the smallest possible order $t = 1$, see [11, Thm 5.2];*
2. *convex quartic bivariate polynomials; here the Lasserre hierarchy is exact at the smallest possible order $t = 2$, see [13, Ex. 3].*

One may prove both these results without using Scheiderer's positivstellensatz (Proposition 3.1), as one can use instead the well-known fact that any nonnegative quadratic polynomial is a sum of squares as well as any nonnegative quartic bivariate polynomial (a result by Hilbert, see e.g. [14]). See also Remark 4.1 below for a further discussion of stronger finite convergence results in special cases.

4. Complexity results. A natural question is whether it is possible to give a bound on the (finite) number of steps required for convergence by the Lasserre hierarchy for problem (1.1) under the assumptions of Corollary 3.3.

Before addressing this question, we briefly discuss known complexity results for convex polynomial optimization, in order to place the discussion in the correct context.

4.1. Recognizing convex problems. A first point to make is that it is NP-hard in the Turing model of computation (described in e.g. [7]) to decide if a given instance of problem (1.1) is a convex optimization problem, due to the following result.

THEOREM 4.1 (Ahmadi et al. [1]). *It is strongly NP-hard in the Turing model of computation to decide if a given form of even degree $d \geq 4$ is convex.*

---

[1] This also follows from Euler's identity: $x^T \nabla^2 f(x) x = d(d-1)f(x)$ for a form f of degree d.

[2] It is interesting to note that Blekherman's proof is not constructive, and no actual examples are known of convex forms that are not sums of squares.

4.2. Complexity results via the ellipsoid method. The best known complexity result for solving problem (1.1) under the assumptions of Corollary 3.3 is by using the ellipsoid method of Yudin-Nemirovski. For given $\epsilon > 0$, the ellipsoid algorithm can compute an ϵ-feasible[3] x such that $|p_0(x) - p_{\min}| \leq \epsilon$ in at most

$$O\left(n^2 \ln\left(\frac{R}{\epsilon}\right)\right)$$

iterations, where each iteration requires the evaluation of the polynomials $p_0, \ldots, p_m$ as well as the gradient of p_0 and of one polynomial that is negative at the current iterate (in order to obtain a separating hyperplane); see e.g. [3, §5.2].

It will be convenient to only consider the real number model (also known as BSS model) of computation [6]. In the real number model, the input is a finite set of real numbers, and an arithmetic operation between two real numbers requires one unit of time. Thus, the size of the input of problem (1.1) may be expressed by four numbers:

1. n, the number of variables;
2. m, the number of constraints;
3. d, the largest total degree of $p_0, \ldots, p_m$;
4. the total number of nonzero coefficients of the polynomials $p_0, \ldots, p_m$ in the standard monomial basis, say $L := \sum_{i=0}^m L_i$, where L_i is the number of nonzero coefficients of p_i.

Note that

$$m + 1 \leq L \leq (m+1)\binom{n+d}{d},$$

and the exact value of L depends on the sparsity of the polynomials $p_0, \ldots, p_m$.

The n-variate polynomial p_i of total degree at most d may be evaluated in at most $O(dL_i)$ arithmetic operations. Thus the total complexity of the ellipsoid method becomes

$$O\left(dLn^2 \ln\left(\frac{R}{\epsilon}\right)\right). \tag{4.1}$$

Note that the ellipsoid algorithm uses the parameter R (and not only the fact that it is finite).

Also, neither the Slater assumption, nor the assumption that the Hessian of the objective is positive definite at a minimizer, is required by the ellipsoid method.

Finally, note that the number of constraints m only enters the complexity bound (4.1) implicitly, via the value L.

4.3. The rank of the Lasserre hierarchy. We now return to the question of giving a bound on the (finite) number of steps required for convergence of the Lasserre hierarchy for problem (1.1) under the assumptions of Corollary 3.3.

Recall that the Lasserre hierarchy computes the values ρ_t in (1.5) as the optimal value of suitable semidefinite programs. The size of the semidefinite program that yields ρ_t is as follows: it has $m+1$ positive semidefinite matrix variables of order $\binom{n+t}{t}$, and there are $\binom{n+2t}{2t}$ linear equality constraints; see [11] or the survey [14] for details on the semidefinite programming reformulations.

[3]We call x ϵ-feasible for problem (1.1) if the ball of radius ϵ and centered at x intersects the feasible set $\mathcal{F}$.

In particular, ρ_t may be computed to ϵ relative accuracy in at most

$$O\left(\left((m+1)\binom{n+2t}{2t}\right)^{\frac{9}{2}} \ln\left(\frac{1}{\epsilon}\right)\right)$$

arithmetic operations using interior point algorithms; see e.g. [3, §6.6.3].

We will call the smallest value of t such that $\rho_t = p_{\min}$ (see (1.1)), the *rank* of the Lasserre hierarchy.

We now show that, in a well-defined sense, the rank of the Lasserre hierarchy must depend on more than just the input size (n, m, d, L) of problem (1.1).

THEOREM 4.2. *Consider problem (1.1) under the assumptions of Corollary 3.3. If deg(p_0) ≥ 4, there is no integer constant $C > 0$ depending only on (n, m, d, L), such that the Lasserre hierarchy converges in C steps.*

Proof. The proof uses a similar construction as in Remark 3.3. As in Remark 3.3, let p be a convex, n-variate form of degree d that is not a sum of squares.

We consider the behavior of the Lasserre hierarchy for the sequence of problems:

$$\min_{x\in\mathbb{R}^n} \left\{ p(x) + \frac{1}{k}\|x\|^2 \;\middle|\; p_1(x) := 1 - \|x\|^2 \geq 0 \right\} \qquad \text{for } k = 1, 2, \ldots \tag{4.2}$$

By construction, for each k, problem (4.2) meets the assumptions of Corollary 3.3. By Corollary 3.3, the Lasserre hierarchy therefore converges in finitely many steps for problem (4.2) for each $k = 1, 2, \ldots$.

Assume now that there exists an integer $t > 0$ such that

$$p + \frac{1}{k}\|x\|^2 \in \mathcal{M}_t(p_1) \quad \forall\, k,$$

where $\mathcal{M}_t(p_1)$ is the truncated quadratic module of degree $2t$ generated by p_1 (see (1.4)). As the set $\{x : p_1(x) \geq 0\}$ has a nonempty interior, the set $\mathcal{M}_t(p_1)$ is closed (see [18] or [14, §3.8]). As a consequence, the limit p of the sequence $p + \frac{1}{k}\|x\|^2$ (as k tends to ∞) must also belong to $\mathcal{M}_t(p_1)$. As explained in Remark 3.3, this contradicts the assumption that p is not a sum of squares. ☐

In the construction used in the proof of Theorem 4.2, the smallest eigenvalue of the Hessian of the objective function in (4.2) at the minimizer $x^* = 0$ tends to zero as $k \to \infty$. This suggests that the rank of the Lasserre hierarchy may depend on the value of the smallest eigenvalue of the Hessian at the minimizer x^*. The smallest eigenvalue of the Hessian at x^* may in turn be viewed as a 'condition number' of the problem that is independent of (n, m, d, L).

REMARK 4.1. *Lasserre [12] showed that the Lasserre rank is bounded by the largest total degree of $p_0, p_1, \ldots, p_m$, if $p_0, -p_1, \ldots, -p_m$ are so-called SOS (sums-of-squares) convex polynomials. More precisely, $p_0 - p_{\min}$ has a decomposition $\sigma_0 + \sum_{j=1}^m \lambda_j p_j$ where σ_0 is a sum of squares and λ_j are nonnegative scalars.*

DEFINITION 4.3. *A polynomial p is called SOS convex if*

$$z^T \nabla^2 p(x) z \text{ is a sum of squares in } (x, z).$$

Ahmadi and Parillo [2] have shown that the SOS convex polynomials form a proper subset of the convex polynomials, and Helton and Nie [8] have shown that SOS convex forms are sums of squares. These results cast some light on our construction in the *proof of Theorem 4.2: the form p used there is not SOS convex, since it is not a sum of squares. Moreover, every 'bad' example like the one in the proof of Theorem 4.2 will necessarily involve convex polynomials that are not SOS convex.*

5. Conclusion and summary. We have given a new proof of the finite convergence of the Lasserre hierarchy for polynomial optimization problems, under weaker assumptions than were known before (Theorem 3.2).

We have also looked at the possibility of bounding the rank of the finite convergence in the convex case, and gave a negative result about the dependence of such a bound on the problem data. In particular, we showed that the number of steps needed for convergence cannot be bounded by a quantity that depends only on the input size (in the real number model of computation). Thus, the worst-case complexity of the ellipsoid method is in some sense better than that of the Lasserre hierarchy for convex problems. Having said that, it is important to remember that the number of operations required by the ellipsoid method will typically equal the worst-case bound, whereas the Lasserre hierarchy can converge quickly for some convex problems (as we discussed in Remarks 3.4 and 4.1). Moreover, the worst-case complexity bound for the Lasserre hierarchy could possibly be improved by deriving error bounds on $p_{\min} - \rho_t$ (see (1.5)) in terms of t. For general polynomial optimization problems, deriving explicit error bounds for the Lasserre hierarchy has proved difficult so far (see [15, 16]), but the additional convexity assumption may simplify this analysis.

Acknowledgements. The authors would like to thank Amir Ali Ahmadi and Marianna Nagy for their valuable comments, and Jean Lasserre for remarks that led to a much improved presentation.

REFERENCES

[1] A. A. Ahmadi, A. Olshevsky, P. A. Parrilo, J. N. Tsitsiklis. NP-hardness of deciding convexity of quartic polynomials and related problems. Preprint, MIT, Cambridge , MA, 2010.

[2] A.A. Ahmadi, P.A. Parrilo. A convex polynomial that is not sos-convex. *Mathematical Programming* (to appear).

[3] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization.* MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.

[4] D.P. Bertsekas. *Nonlinear Programming: 2nd Edition*, Athena Scientific, 2004.

[5] G. Blekherman. Convex forms that are not sums of squares. arXiv:0910.0656v1, October 2009.

[6] L. Blum, M. Shub, and S. Smale: On a theory of computation and complexity over the real numbers: NP-completeness, recursive functions and universal machines. *Bull. Amer. Math. Soc. (N.S.)* **21**(1) (1989), 1–46.

[7] M.R. Garey and D.S. Johnson. *Computers and intractability: a guide to the theory of NP–completeness.* W.H. Freeman and Company, Publishers, San Fransisco, USA, 1979.

[8] J. W. Helton and J. Nie. Semidefinite representation of convex sets, *Math. Program.*, 122(1):21–64, 2010.

[9] V. Jeyakumar. Farkas lemma: Generalizations. in *Encyclopedia of Optimization II*, C.A. Floudas and P. Pardalos Eds., Kluwer, Dordrecht, 87–91, 2001.

[10] E. Klerk, M. Laurent, P. Parrilo. On the equivalence of algebraic approaches to the minimization of forms on the simplex. In D. Henrion and A. Garulli (Eds.), *Positive Polynomials in Control*. Berlin: Springer Verlag. (LNCIS, 312), 2005.

[11] J.B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, **11**:296–817, 2001.

[12] J.B. Lasserre. Convexity in semialgebraic geometry and polynomial optimization. *SIAM Journal on Optimization* **19**, 1995–2014, 2009.

[13] J.B. Lasserre. Convex sets with semidefinite representation. *Math. Program., Ser. A*,120:457–477, 2009.

[14] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerg-*

ing Applications of Algebraic Geometry, Vol. 149 of IMA Volumes in Mathematics and its Applications, M. Putinar and S. Sullivant (eds.), Springer, pages 157–270, 2009.

[15] J. Nie. An approximation bound analysis for Lasserre's relaxation in multivariate polynomial optimization. Preprint, 2010. `http://math.ucsd.edu/~njw/PUBLICPAPERS/LasBd.pdf`

[16] J. Nie and M. Schweighofer. On the complexity of Putinar's Positivstellensatz. *Journal of Complexity*, 23(1):135–150, 2007.

[17] J. Nocedal and S. Wright. *Numerical Optimization*, Springer, 2000.

[18] V. Powers and C. Scheiderer. The moment problem for non-compact semialgebraic sets. *Adv. Geom.*, 1:71–88, 2001.

[19] C. Scheiderer. Sums of squares on real algebraic curves. *Mathematische Zeitschrift*, **245**, 725-760, 2003.